Panel Data and Tracking

Damir Esenaliev, SIPRI/ISDC

Measuring Violent Conflict in Micro-Level Surveys
March 19-23, 2018
Perugia, Italy
Overview

1. Features of panel data
2. Tracking principles
3. Young Lives Study
4. Life in Kyrgyzstan Study
1. Features of panel data
Panel Data

- **Panel Data** is data in which we observe repeated cross-sections of the same individuals, households, or companies.

- Examples:
  - Annual unemployment rates of each state over several years
  - Quarterly sales of firms over several quarters
  - Wages for the same worker, working at several different jobs
The Case for an Individual Panel

Repeated Cross-Section

Panel

Tracking the same people enables causal analysis!
Panel Data Advantages

• The key feature of panel data is that we observe the same individual in more than one condition.
  – not so much “Who is unemployed?”, but rather “Will people who are unemployed ever find a job?”

• Panel data is better suited for studying the dynamics of change and understanding transition behavior, for example:
  – migration of a household member
  – company bankruptcy or merger
  – technological change or economic shocks

• Omitted variables that are fixed will take on the same values each time we observe the same individual.
  – A person’s upbringing, family characteristics, innate ability, and demographics (except age) do not change.
Panel Data Advantages-2

• Capturing rare life events
  – getting (grand) children, choosing a mortgage, being a victim of a crime.

• Practical considerations
  – Use a sub-sample of panel for focusing on certain topics: ask about mortgage
  – only HHs who own houses

• Methodological benefits
  – Test measurement instruments for reliability over time.
Example: Poverty dynamics & classes of HHs

Rich HHs

Transitory HHs

Poverty line

Income

T1 T2 T3 T4 T5
Types of Longitudinal Study

1. **Trend Studies** – these examine change within a population over time (e.g. the Census).

2. **Cohort Studies** – these examine over time specific subpopulations or cohorts (often, although not necessarily, the same individuals) e.g. a study might interview people aged 30 in 1970, 40 in 1980, 50 in 1990 and 60 in 2000.

3. **Panel Study** – These examine the same set of people each time.
Rotating panel model

<table>
<thead>
<tr>
<th></th>
<th>t₁</th>
<th>t₂</th>
<th>t₃</th>
<th>t₄</th>
<th>t₅</th>
<th>t₆</th>
<th>t₇</th>
</tr>
</thead>
<tbody>
<tr>
<td>g₁</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₄</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₅</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₆</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₇</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g₈</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Different types of panel are, for example:

1. Volunteer opt-in panels (Harris Interactive, Greenfield, NFO)
2. Pre-recruited panels of Internet users (Pew Research Center)
3. Pre-recruited panels of full population (Knowledge Networks, CentERdata)
4. Direct access panels, Internet access panels
5. Special purpose panels: scatter panel, people-meter panel (AC Nielsen)
6. Rotating panels
7. Cohort studies

25.4 HOW TO SET UP AND MAINTAIN A PANEL SURVEY

Already in the planning stages, it is very important to realise that a panel needs maintenance. In fact, you will need some effective panel management tools to prevent you from drowning in management problems (if not in data). The panel management system becomes more critical as the time between waves gets shorter. Setting up a panel survey is basically not much different from setting up a cross-sectional survey, except that you need to keep the address of the respondent for a future contact. However, as soon as you collected the data in the first wave, the time until the next wave is ticking away, and you had better be prepared. Maintaining a panel survey requires that you are in control of the information flow and that you are able to react adequately if necessary.

To illustrate that controlling the information flow in a panel survey is not straightforward consider the simple question, “How many respondents are in the panel?” or “What is the response rate?” in the case of an Internet access
## Existing panel surveys

<table>
<thead>
<tr>
<th>Panel Survey</th>
<th>Sample of HHs</th>
<th>Sample of individuals</th>
<th>Started</th>
<th># of waves</th>
</tr>
</thead>
<tbody>
<tr>
<td>British Household Panel Survey, BHPS</td>
<td>5,500</td>
<td>10,300</td>
<td>1991</td>
<td>many</td>
</tr>
<tr>
<td>German Socio-Economic Panel, GSOEP</td>
<td>11,000</td>
<td>30,000</td>
<td>1984</td>
<td>many</td>
</tr>
<tr>
<td>The Life in Kyrgyzstan Panel Study, LiK</td>
<td>3,000</td>
<td>8,160</td>
<td>2010</td>
<td>4</td>
</tr>
<tr>
<td>Russian Longitudinal Measurement Survey, RLMS</td>
<td>6,334</td>
<td>17,154</td>
<td>1992</td>
<td>23</td>
</tr>
<tr>
<td>Ukrainian Longitudinal Monitoring Survey, URLMS</td>
<td>4,056</td>
<td>7,200</td>
<td>2003</td>
<td>3</td>
</tr>
<tr>
<td>The Panel Survey of Income Dynamics, USA</td>
<td>5,000</td>
<td>18,000</td>
<td>1968</td>
<td>many</td>
</tr>
<tr>
<td>Young Lives Study of childhood poverty in Ethiopia, India, Peru and Vietnam</td>
<td>±12,000</td>
<td>12,000</td>
<td>2002</td>
<td>4</td>
</tr>
</tbody>
</table>
2. Tracking in panel data
Attrition in panel surveys

- Sample attrition – loss of a respondent in a panel - is one of the challenges for panel surveys due to
  - Migration
  - Refusal
  - Absence
  - Death

- The number of respondents who do not participate in each round of data collection is likely to cumulate over time.

- Attrition makes the initial panel sample less representative as the attrition can be non-random

- Thus, preventing high attrition maximizes the precision of an analysis based on panel data and strengthens its statistical validity.
Tracking in panel surveys

• Tracking respondents is to follow respondents who moved from the location where they were interviewed the last time.
• Tracking individuals and households can entail significant costs and may require specific focus by the organization conducting the survey.
• The survey planners make decisions
  – what is the panel unit: dwelling, household, individual.
  – about whether to follow households and individuals who left their baseline location
  – to what degree with respect to time, distance, and cost.
Panel sample maintenance system

- PANEL NEEDS MAINTENANCE
- Has a dedicated tracking team
- Keeps addresses and contact information of respondents
  - Not only from a respondent, but also from two additional contacts of people who may know about location of a respondent
- Ensures the frequent contact with the respondents and builds rapport
- Has the protocols to follow the movers
- Has the optional ways of providing data from the respondents who refuse
- Has information to show to respondents about usefulness of the data and underlying research
3. Tracking in Young Lives Study
Features of the study

- Young Lives is an international study launched in 2002 by Oxford University of childhood poverty following 12,000 children in Ethiopia, India, Peru and Vietnam over 15 years.
Tracking system in Young Lives Study

- Tracking system is established that updates regularly the location and contact information of a child.
- Wherever possible, researchers trace the new location of children who have moved and visit them at their new address.
- If a child leaves the country, the efforts made to connect to him/her by phone or online.
- Peru has higher attrition rates because of migration for work or study reasons ...
- ... while in Vietnam the attrition is lower as movements are restricted.
Attrition in the Young Lives Study

Table 1. Attrition rates by category and country, Rounds 1 to 2 (Outes Leon and Dercon 2008: 5)

<table>
<thead>
<tr>
<th>Country</th>
<th>Sample Size</th>
<th>Child died</th>
<th>Refused to answer</th>
<th>Untraceable</th>
<th>Attrition rate (including deaths)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethiopia</td>
<td>2,998</td>
<td>67</td>
<td>11</td>
<td>31</td>
<td>1.43</td>
</tr>
<tr>
<td>India</td>
<td>3,019</td>
<td>35</td>
<td>14</td>
<td>25</td>
<td>1.31</td>
</tr>
<tr>
<td>Peru</td>
<td>2,766</td>
<td>6</td>
<td>64</td>
<td>33</td>
<td>3.51</td>
</tr>
<tr>
<td>Vietnam</td>
<td>3,000</td>
<td>13</td>
<td>3</td>
<td>16</td>
<td>0.64</td>
</tr>
<tr>
<td>Total</td>
<td>11,783</td>
<td>121</td>
<td>92</td>
<td>105</td>
<td>2.7</td>
</tr>
</tbody>
</table>
Some lessons from Young Lives Study

• Minimizing attrition rates is about more than being able to locate children.

• Equally important is talking to respondents who want to drop out to understand their reasons and keep them by addressing their concerns.

• Some respondents are unhappy about the length and complexity of the survey, and the amount of time needed to complete it.

• Ensuring that the survey is well-paced and contains different methods for engaging respondents is important for keeping attrition low.
4. Life in Kyrgyzstan Panel Study
LiK panel design

• The LiK is an individual panel, not a household panel.
• All adult members of the households, not just one respondent, are interviewed and tracked over time.
• In principle, all persons who took part in the first wave of the survey in 2010 are to be surveyed in the following waves.
• If a individual moves within Kyrgyzstan, the individual is followed (though, not students); if he/she moves abroad, the individual is dropped from the sample.
• In a case of in-country move, other members (e.g. spouse and children) of the new household are then included in the sample as well.
LiK sample attrition

Table 1: Sample evolution of the Life in Kyrgyzstania Survey

<table>
<thead>
<tr>
<th>Wave</th>
<th># of HHs</th>
<th># of individuals</th>
<th># of households as % to 2010</th>
<th># of individuals as % to 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>3,000</td>
<td>8,160</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2011</td>
<td>2,862</td>
<td>8,065</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>2012</td>
<td>2,816</td>
<td>8,177</td>
<td>94</td>
<td>100</td>
</tr>
<tr>
<td>2013</td>
<td>2,586</td>
<td>7,681</td>
<td>86</td>
<td>94</td>
</tr>
<tr>
<td>2016</td>
<td>2,529</td>
<td>8,094</td>
<td>84</td>
<td>99</td>
</tr>
</tbody>
</table>
Lessons from LiK Study

• Has been quite successful in retaining the sample of individuals
• No dedicated tracking system between waves, but the tracking protocols are embedded into the field work cycles
• Main reason for attrition is migration abroad
• Difficult to persuade the movers to stay, as in most cases the move was related to unfortunate life events
Summary

- Panel data – when we observe the same objects over time - have beneficial features compared to cross-sectional data

- Attrition is an inevitable feature of panel data

- The tracking efforts are important to keep the representativeness of the panel sample

- Many panel studies do successfully maintain their sample, but at the cost of greater dedicated efforts and costs

- In-conflict and post-conflict areas represent a greater challenge for panel studies
References to panel studies

• Firman Witoelar, Tracking in Longitudinal Household Surveys, WB LSMS-ISA Paper, 2011.
• The British Household Panel Survey; https://www.iser.essex.ac.uk/bhps
• Ukrainian Longitudinal Monitoring Survey; https://datasets.iza.org/dataset/56/ukrainian-longitudinal-monitoring-survey
• The Panel Study of Income Dynamics (PSID), USA; https://psidonline.isr.umich.edu/default.aspx
• Young Lives Study; http://younglives.qeh.ox.ac.uk/.